



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Choulika, Andre
Perrin, Arnaud
Dujon, Bernard
Nicolas, Jean-Francois
- (ii) TITLE OF INVENTION: Nucleotide Sequence Encoding the Enzyme
I-SCEI and the Uses Thereof
- (iii) NUMBER OF SEQUENCES: 52
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunner
 - (B) STREET: 1300 I Street, N.W.
 - (C) CITY: Washington
 - (D) STATE: D.C.
 - (E) COUNTRY: USA
 - (F) ZIP: 20005-3315
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 09/244,130
 - (B) FILING DATE: 04-FEB-1999
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 09/119,024
 - (B) FILING DATE: 20-JUN-1998
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/336,241
 - (B) FILING DATE: 07-NOV-1994
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/971,160
 - (B) FILING DATE: 05-NOV-1992
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/879,689
 - (B) FILING DATE: 05-MAY-1992
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Meyers, Kenneth J.
 - (B) REGISTRATION NUMBER: 25,146
 - (C) REFERENCE/DOCKET NUMBER: 3495-0111-10
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 202-408-4000
 - (B) TELEFAX: 202-408-4400

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 714 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGCATATGA AAAACATCAA AAAAAACCAG GTAATGAACC TCGGTCCGAA CTCTAAACTG	60
CTGAAAGAAT ACAAATCCCA GCTGATCGAA CTGAACATCG AACAGTTCGA AGCAGGTATC	120
GGTCTGATCC TGGGTGATGC TTACATCCGT TCTCGTGATG AAGGTAAAAC CTACTGTATG	180
CAGTTCGAGT GGAAAAACAA AGCATACATG GACCACGTAT GTCTGCTGTA CGATCAGTGG	240
GTACTGTCCC CGCCGCACAA AAAAGAACGT GTTAACCACC TGGGTAACCT GGTAATCACC	300
TGGGGCGCCC AGACTTTCAA ACACCAAGCT TTCAACAAAC TGGCTAACCT GTTCATCGTT	360
AACAACAAAA AAACCATCCC GAACAACCTG GTTGAAAACCT ACCTGACCCC GATGTCTCTG	420
GCATACTGGT TCATGGATGA TGGTGGTAAA TGGGATTACA ACAAAAACCTC TACCAACAAA	480
TCGATCGTAC TGAACACCCA GTCTTTCACT TTCGAAGAAG TAGAATACCT GGTTAAGGGT	540
CTGCGTAACA AATTCCAACCT GAACTGTTAC GTAAAAATCA ACAAAAACAA ACCGATCATC	600
TACATCGATT CTATGTCTTA CCTGATCTTC TACAACCTGA TCAAACCGTA CCTGATCCCG	660
CAGATGATGT ACAAACCTGCC GAACACTATC TCCTCCGAAA CTTTCCTGAA ATAA	714

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 237 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	His	Met	Lys	Asn	Ile	Lys	Lys	Asn	Gln	Val	Met	Asn	Leu	Gly	Pro
1				5				10						15	
Asn	Ser	Lys	Leu	Leu	Lys	Glu	Tyr	Lys	Ser	Gln	Leu	Ile	Glu	Leu	Asn
			20					25					30		

Ile Glu Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr
 35 40 45
 Ile Arg Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp
 50 55 60
 Lys Asn Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp
 65 70 75 80
 Val Leu Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn
 85 90 95
 Leu Val Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn
 100 105 110
 Lys Leu Ala Asn Leu Phe Ile Val Asn Asn Lys Lys Thr Ile Pro Asn
 115 120 125
 Asn Leu Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe
 130 135 140
 Met Asp Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys
 145 150 155 160
 Ser Ile Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr
 165 170 175
 Leu Val Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys
 180 185 190
 Ile Asn Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu
 195 200 205
 Ile Phe Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr
 210 215 220
 Lys Leu Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys
 225 230 235

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 722 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AAAAATAAAA TCATATGAAA AATATTAAAA AAAATCAAGT AATCAATCTC GGTCCTATTT	60
CTAAATTATT AAAAGAATAT AAATCACAAT TAATTGAATT AAATATTGAA CAATTTGAAG	120
CAGGTATTGG TTTAATTTTA GGAGATGCTT ATATTCGTAG TCGTGATGAA GGTAAAACTT	180

ATTGTATGCA ATTTGAGTGG AAAAATAAGG CATACATGGA TCATGTATGT TTATTATATG 240
 ATCAATGGGT ATTATCACCT CCTCATAAAA AAGAAAGAGT TAATCATTTA GGTAATTTAG 300
 TAATTACCTG GGGAGCTCAA ACTTTTAAAC ATCAAGCTTT TAATAAATTA GCTAACTTAT 360
 TTATTGTAAA TAATAAAAAA CTTATTCCTA ATAATTTAGT TGAAAATTAT TTAACACCTA 420
 TGAGTCTGGC ATATTGGTTT ATGGATGATG GAGGTAAATG GGATTATAAT AAAAATTCTC 480
 TTAATAAAAG TATTGTATTA AATACACAAA GTTTTACTTT TGAAGAAGTA GAATATTTAC 540
 TTAAAGGTTT AAGAAATAAA TTTCAATTAA ATTGTTATGT TAAAATTAAT AAAAATAAAC 600
 CAATTATTTA TATTGATTCT ATGAGTTATC TGATTTTSTA TAATTTAATT AAACCTTATT 660
 TAATTCCTCA AATGATGTAT AACTGCCTA ATACTATTTC ATCCGAAACT TTTTAAAT 720
 AA 722

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 235 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Lys	Asn	Ile	Lys	Lys	Asn	Gln	Val	Met	Asn	Leu	Gly	Pro	Asn	Ser	1	5	10	15
Lys	Leu	Leu	Lys	Glu	Tyr	Lys	Ser	Gln	Leu	Ile	Glu	Leu	Asn	Ile	Glu	20	25	30	
Gln	Phe	Glu	Ala	Gly	Ile	Gly	Leu	Ile	Leu	Gly	Asp	Ala	Tyr	Ile	Arg	35	40	45	
Ser	Arg	Asp	Glu	Gly	Lys	Thr	Tyr	Cys	Met	Gln	Phe	Glu	Trp	Lys	Asn	50	55	60	
Lys	Ala	Tyr	Met	Asp	His	Val	Cys	Leu	Leu	Tyr	Asp	Gln	Trp	Val	Leu	65	70	75	80
Ser	Pro	Pro	His	Lys	Lys	Glu	Arg	Val	Asn	His	Leu	Gly	Asn	Leu	Val	85	90	95	
Ile	Thr	Trp	Gly	Ala	Gln	Thr	Phe	Lys	His	Gln	Ala	Phe	Asn	Lys	Leu	100	105	110	
Ala	Asn	Leu	Phe	Ile	Val	Asn	Asn	Lys	Lys	Leu	Ile	Pro	Asn	Asn	Leu	115	120	125	
Val	Glu	Asn	Tyr	Leu	Thr	Pro	Met	Ser	Leu	Ala	Tyr	Trp	Phe	Met	Asp				

130	135	140
Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Leu Asn Lys Ser Ile		
145	150	155
Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Cys Tyr Leu Val		
	165	170
Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys Ile Asn		
	180	185
Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu Ile Phe		
	195	200
Tyr Asn Ile Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr Lys Leu		
	210	215
Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys		
225	230	235

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 754 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCGGATCCAT GCATATGAAA AACATCAAAA AAAACCAGGT AATGAACCTG GGTCCGAAC	60
CTAAACTGCT GAAAGAATAC AAATCCCAGC TGATCGAACT GAACATCGAA CAGTTCGAAG	120
CAGGTATCGG TCTGATCCTG GGTGATGCTT ACATCCGTTT TCGTGATGAA GGTAAACCT	180
ACTGTATGCA GTTCGAGTGG AAAAACAAAG CATACTGGA CCACGTATGT CTGCTGTACG	240
ATCAGTGGGT ACTGTCCCCG CCGCACAAAA AACAACGTGT TAACCACCTG GGTAACCTGG	300
TAATCACCTG GGGCGCCCAG ACTTTCAAAC ACCAAGCTTT CAACAACTG GCTAACCTGT	360
TCATCGTTAA CAACAAAAAA ACCATCCCGA ACAACCTGGT TGAAACTAC CTGACCCCGA	420
TGTCTCTGGC ATACTGGTTC ATGGATGATG GTGGTAAATG GGATTACAAC AAAAACTCTA	480
CCAACAAATC GATCGTACTG AACACCCAGT CTTTCACTTT CGAAGAAGTA GAATACCTGG	540
TTAAGGGTCT GCGTAACAAA TTCCAAGTGA ACTGTTACGT AAAAATCAAC AAAACAAAC	600
CGATCATCTA CATCGATTCT ATGTCTTACC TGATCTTCTA CAACCTGATC AAACCGTACC	660
TGATCCCGCA GATGATGTAC AACTGCCGA AACTATCTC CTCCGAACT TTCCTGAAAT	720

AATAAGTCGA CTGCAGGATC CGGTAAGTAA GTAA

754

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 11 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AATGCTTCC A

11

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTTACGCTAG GGATAACAGG GTAAT

25

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CAATGCGATC CCTATTGTCC CATTA

25

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1738 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG GGAGCTTCCA	60
GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG ACTTGAGCGT	120
CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG CAACGCGGCC	180
TTTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTTCC TCGGTTATCC	240
CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC TCGCCGCAGC	300
CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC AATACGCAAA	360
CCGCCTCTCC CCGCGCGTTG GCCGATTCAT TAATGCAGCT GGCACGACAG GTTTCCCGAC	420
TGGAAAGCGG GCAGTGAGCG CAACGCAATT AATGTGAGTT AGCTCACTCA TTAGGCACCC	480
CAGGCTTTAC ACTTTATGCT TCCGGCTCGT ATGTTGTGTG GAATTGTGAG CGGATAACAA	540
TTTCACACAG GAAACAGCTA TGACCATGAT TACGAATTCT CATGTTTGAC AGCTTATCAT	600
CGATAAGCTT TAATGCGGTA GTTTATCACA GTTAAATTGC TAACGCAGTC AGGCACCGTG	660
TATGAAATCT AACAATGCGC TCATCGTCAT CCTCGGCACC GTCACCCTGG ATGCTGTAGG	720
CATAGGCTTG GTTATGCCGG TACTGCCGGG CCTCTTGCGG GATATCCGCC TGATGCGTGA	780
ACGTGACGGA CGTAACCACC GCGACATGTG TGTGCTGTTC CGCTGGGCAT GCCAGGACAA	840
CTTCTGGTCC GGTAACGTGC TGAGCCCGGC CAAGCTTACT CCCCATCCCC CTGTTGACAA	900
TTAATCATCG GCTCGTATAA TGTGTGGAAT TGTGAGCGGA TAACAATTTT ACACAGGAAA	960
CAGGATCCAT GCATATGAAA AACATCAAAA AAAACCAGGT AATGAACCTG GGTCCGAAC	1020
CTAAACTGCT GAAAGAATAC AAATCCCAGC TGATCGAACT GAACATCGAA CAGTTCGAAG	1080
CAGGTATCGG TCTGATCCTG GGTGATGCTT ACATCCGTTC TCGTGATGAA GGTA AACCT	1140
ACTGTATGCA GTTCGAGTGG AAAAACAAAG CATACTGGA CCACGTATGT CTGCTGTACG	1200
ATCAGTGGGT ACTGTCCCCG CCGCACAAAA AAGAACGTGT TAACCACCTG GGTAACCTGG	1260
TAATCACCTG GGGCGCCCAG ACTTTCAAAC ACCAAGCTTT CAACAACTG GCTAACCTGT	1320
TCATCGTTAA CAACAAAAAA ACCATCCCGA ACAACCTGGT TGAAACTAC CTGACCCCGA	1380
TGTCTCTGGC ATACTGGTTC ATGGATGATG GTGGTAAATG GGATTACAAC AAAA ACTCTA	1440
CCAACAAATC GATCGTACTG AACACCCAGT CTTTCACTTT CGAAGAAGTA GAATACCTGG	1500
TTAAGGGTCT GCGTAACAAA TTCCA ACTGA ACTGTTACGT AAAAATCAAC AAAAACA AAC	1560

CGATCATCTA CATCGATTCT ATGTCTTACC TGATCTTCTA CAACCTGATC AAACCGTACC 1620
 TCATCCCCCA GATGATGTAC AAAGTCCGA AACTATCTC CTCCGAAACT TTCCTGAAAT 1680
 AATAAGTCGA CCTGCAGCCC AAGCTTGGCA CTGGCCGTCG TTTTACAACG TCGTGACT 1738

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met	Leu	Val	Arg	Gly	Ala	Glu	Pro	Met	Glu	Lys	Arg	Gln	Gln	Arg	Gly
1				5				10						15	
Leu	Phe	Thr	Val	Pro	Gly	Leu	Leu	Leu	Ala	Phe	Cys	Ser	His	Val	Leu
			20				25						30		
Ser	Cys	Val	Ile	Pro											
			35												

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met	Gln	Leu	Ala	Arg	Gln	Val	Ser	Arg	Leu	Glu	Ser	Gly	Gln
1			5					10					

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 13 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Leu Pro Ala Arg Met Leu Cys Gly Ile Val Ser Gly

1

5

10

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 9 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Thr Met Ile Thr Asn Ser His Val
 1 5

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 80 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Lys Ser Asn Asn Ala Leu Ile Val Ile Leu Gly Thr Val Thr Leu
 1 5 10 15

Asp Ala Val Gly Ile Gly Leu Val Met Pro Val Leu Pro Gly Leu Leu
 20 25 30

Arg Asp Ile Arg Leu Met Arg Glu Arg Asp Gly Arg Asn His Arg Asp
 35 40 45

Met Cys Val Leu Phe Arg Trp Ala Cys Gln Asp Asn Phe Trp Ser Gly
 50 55 60

Asn Val Leu Ser Pro Ala Lys Leu Thr Pro His Pro Pro Val Asp Asn
 65 70 75 80

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 7 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Cys Gly Ile Val Ser Gly
1 5

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 237 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met His Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro
1 5 10 15
Asn Ser Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn
20 25 30
Ile Glu Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr
35 40 45
Ile Arg Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp
50 55 60
Lys Asn Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp
65 70 75 80
Val Leu Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn
85 90 95
Leu Val Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn
100 105 110
Lys Leu Ala Asn Leu Phe Ile Val Asn Asn Lys Lys Thr Ile Pro Asn
115 120 125
Asn Leu Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe
130 135 140
Met Asp Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys
145 150 155 160
Ser Ile Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr
165 170 175
Leu Val Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys
180 185 190
Ile Asn Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu
195 200 205
Ile Phe Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr

210

215

220

Lys Leu Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys
225 230 235

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGCTAGGGAT AACAGGGTAA TATAGC

26

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCGATCCCTA TTGTCCCATT ATATCG

26

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TTCTCATGAT TAGCTCTAAT CCATGG

26

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

AAGAGTACTA ATCGAGATTA GGTACC

26

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTTTGGTCAT CCAGAAGTAT ATATTT

26

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GAAACAGTA GGTCTTCATA TATAAA

26

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TAACGGTCCT AAGGTAGCGA AATTCA

26

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ATTGCCAGGA TTCCATCGCT TTAAGT

26

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TGACTCTCTT AAGGTAGCCA AATGCC

26

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ACTGAGAGAA TTCCATCGGT TTACGG

26

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CGAGGTTTTG GTAAC TATT ATTACC

26

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CCTCCAAAAC CATTGATAAA TAATGG

26

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GGGTTCAAAA CGTCGTGAGA CAGTTT

26

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CCCAAGTTTT GCAGCACTCT GTCAAA

26

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GATGCTGTAG GCATAGGCTT GGTTAT

26

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTACGACATC CGTATCCGAA CCAATA

26

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTTTCCGCAA CAGTATAATT TTATAA

26

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GAAAGGCGTT GTCATATTAA AATATT

26

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

ACCATGGGGT CAAATGTCTT TCTGGG

26

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TGGTACCCCA GTTTACAGAA AGACCC

26

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GTGCCTGAAT GATATTTATT ACCTTT

26

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:
GTGCCTGAAT GATATTTATT ACCTTT

26

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 39 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:
CAACGCTCAG TAGATGTTTT CTTGGGTCTA CCGTTTAAT

39

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 39 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:
GTTGCGAGTC ATCTACAAAA GAACCCAGAT GGCAAATTA

39

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:
CAAGCTTATG AGTATGAAGT GAACACGTTA TT

32

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GTTCTGAATAC TCATACTTCA CTTGTGCAAT AA

32

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GCTATTTCGTT TTTATGTATC TTTTGCGTGT AGCTTTAA

38

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

CGATAAGCAA AAATACATAG AAAACGCACA TGGAATT

38

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 80 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CCAAGCTCGA ATTCGCATGC TCTAGAGCTC GGTACCCGGG ATCCTGCAGT CGACGCTAGG

60

GATAACAGGG TAATACAGAT

80

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 80 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GGTTCGAGCT TAAGCGTACG AGATCTCGAG CCATGGGCCC TAGGACGTCA GCTGCGATCC
CTATTGTCCC ATTATGTCTA

60

80

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 80 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

ATCAGATCTA AGCTTGCATG CCTGCAGGTC GACTCTAGAG GATCCCCGGG TACCGAGCTC
GAATTCAGTG GCCGTCGTTT

60

80

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 80 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TAGTCTAGAT TCGAACGTAC GGACGTCCAG CTGAGATCTC CTAGGGGCCC ATGGCTCGAG
CTTAAGTGAC CGGCAGCAAA

60

80

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 80 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

TACAACGTCG TGACTGGGAA AACCTGGCG TTACCCAAC TAATCGCCTT GCAGCACATC 60
CCCCTTTCGC CAGCTGGCGT 80

(2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 80 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

ATGTTGCAGC ACTGACCCTT TTGGGACCGC AATGGGTGA ATTAGCGGAA CGTCGTGTAG 60
GGGGAAAGCG GTCGACCGCA 80

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

TAGGGATAAC AGGGTAAT 18

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

ATCCCTATTG TCCCATTA

18